November, 1964

ON A CLASS OF ADMISSIBLE PARTITIONS

T. Cacoullos

Technical Report No. 52

University of Minnesota
Minneapolis, Minnesota

ON A CLASS OF ADMISSIBLE PARTITIONS

by

T. Cacoullos

University of Minnesota

0. Introduction and Summary. In many multi-decision problems the existence of admissible decision functions (for definitions we refer to [5]) depends upon the existence of corresponding partitions of the sample space into regions of specified shape. Usually the requirements for such statistical partitions differ from those relating to less restricted partitions of a set S according to a vector of finite measures on the measurable subsets of S, usually referred to under the general title of the "ham sandwich problem" (see eg, [3], where further references may be found). Nevertheless, as indicated here the solutions to a wide class of division problems rest very heavily on the fundamental result of Lyapunov and generalizations of this (see, eg. [4])

Several multi-decision problems (Section 4 below) relating to the mean μ of a k-variate normal distribution $N(\mu,\Sigma)$ reduce to the problem of locating (hence called "topothetical", cf.[2]) the parameter point μ into one of $k + 1$ convex k-dimensional polyhedral cones, $\omega_1, \omega_2, \ldots, \omega_{k+1}$ (hereafter referred to as "cones") with common vertex μ_0 which form a partition of the parameter space E_k of μ (see (1) below). Let us identify the sample space E_k of an observation X from $N(\mu,\Sigma)$ with the parameter space E_k. It was shown in [2] that the family $\mathcal{R}_\omega$ of all translations $R(\tau) = (R_1(\tau), \ldots, R_{k+1}(\tau))$ of the system $\omega = (\omega_1,\ldots,\omega_{k+1})$ (see Definition 2) defines a class of admissible procedures, henceforth referred to as partitions; the decision d_i that $\mu \in \omega_i$ is taken when the observation $x \in R_i$. Furthermore, there exists a unique minimax partition $R(\tau_o) \in \mathcal{R}_\omega$. The minimax character of $R(\tau_o)$ amounts to the following proposition: There exists a unique partition of E_k into k+1 cones with the same probability content under the normal k-variate distribution. The distribution may be assumed spherical normal (unit variance in any direction) without any loss of generality, since a nonsingular linear transformation T such that $T \Sigma T' = I$ preserves the shape of the partition ω.

The purpose of this note is to extend the above result to the case of arbitrary probability contents for such conical regions (Theorem 1), and at the same time show how the corresponding partitions are related to classes of admissible partitions for a family of classification and topothetical problems relating to the normal mean μ. Several problems which have been extensively studied in the statistical literature emerge as special cases of our general topothetical problem (Section 4).

1. Preliminary results. For the proof of the main result (Theorem 1), we require certain preliminary results, which are of some interest in themselves. First some notation and definitions.

Let σ denote an arbitrary but fixed k-simplex with vertices $\mu_1, \ldots, \mu_{k+1}$ and center (i.e., the center of the hypersphere passing through the point $\mu_1, \ldots, \mu_{k+1}$) the common vertex μ_o of the cones $\omega_1, \ldots, \omega_{k+1}$, such that for each $i = 1, \ldots, k+1$, if d denotes the usual distance function in E_k,

(1) $\omega_i = \{\mu \in E_k: \quad \delta_{ij}(\mu) \leq 0, \ j \neq i, \ j = 1, \ldots, k+1\}$

where for $i \neq j$

(2) $\delta_{ij}(\mu) \equiv d^2(\mu, \mu_i) - d^2(\mu, \mu_j) = (2\mu - \mu_i - \mu_j)' (\mu_j - \mu_i)$.

Remark. The simplex σ is characterized by the property that the k bounding half-hyperplanes of ω_i are the perpendicular bisectors of the edges of σ; the hyperplane $\delta_{ij}(\mu) = 0$ is perpendicular to the edge (μ_i, μ_j). However, the exposition below shows that any k-simplex whose edges through vertex μ_i are perpendicular to the bounding hyperplanes of ω_i would suffice for our purposes. Note also that convexity of ω_i is essential for the existence of σ above.

Denfinition 1. For any k-simplex σ^* the corresponding classification problem of choosing one of its $k + 1$ vertices as the true mean μ on the basis of an observation X from $N(\mu, I)$ will be called the σ^*-classification problem.

Assumption. Throughout this paper we assume a simple loss function, i.e., 0 or 1 according to whether a correct or incorrect decision is taken. Therefore the risk function becomes the probability of error.

Denfition 2. Any partition R in $\mathcal{R}_\omega$, defined as a translation of the system $\omega = (\omega_1, \ldots, \omega_{k+1})$, will be called a similar partition to ω. The class $\mathcal{R}_\omega$ coincides with the totality of partitions $R = (R_1, \ldots, R_k)$ where

$$R_i = \{x \in E_k: \delta_{ij}(x) \leq c_i - c_j,\ j \neq i\},\ i = 1, \ldots, k + 1, \text{and}$$

$c = (c_1, \ldots, c_{k+1})$ is a constant vector of non-negative components (cf. [1]).

The following two lemmas summarize relevant results obtained in [2].

Lemma 1. The class of conical partitions $\mathcal{R}_\omega$ is

(a) the minimal complete class of partitions (procedures) for the σ-classification problem (Theorem 6.7.1. of [1]),

(b) an admissible class of partitions for the topothetical problem of locating μ into one of the cones $\omega_1, \ldots, \omega_{k+1}$. (Theorem 1.8.4 of [2]).

Lemma 2. The (unique) admissible minimax partition for the σ-classification problem is such

(a) for the topothetical probelm of locating μ into one of the conical regions

$$\omega_i(\sigma) = \{\mu \in E_k: \delta_{ij}(\mu) \leq -d^2(\mu_i, \mu_j),\ j \neq i\},\ i = 1, \ldots, k + 1,$$

with vertices μ_i, respectively, the complement of $\omega_1(\sigma) + \omega_2(\sigma) + \ldots + \omega_{k+1}(\sigma)$ consituting an indifference region. (Theorem 1.7.1 of [2]).

(b) for the topothetical problem of locating μ into one of the subsets

$\omega_1^*, \ldots, \omega_{k+1}^*$ of the $\omega_1(\sigma), \ldots, \omega_{k+1}(\sigma)$, respectively, provided ω_i^* contains the point μ_i and the complement of $\omega_1^* + \omega_2^* + \ldots + \omega_{k+1}^*$ constitutes an indifference region. (Follows immediately from (a) and Lemma 1.7.3 of [2]).

Lemma 3. For each $r \gtrless 0$, denote by σ_r the homothetic k-simplex of σ with center of similitude the point μ_0, vertices $\mu_1(r), \ldots, \mu_{k+1}(r)$ and homothetic ratio r $(|\mu_0 - \mu_i(r)| = |\mu_0 - \mu_i| r)$. Let $p_i(\delta)$ denote the probability of taking decision D_i that $\mu = \mu_i(r)$ in the σ_r-classification problem when using procedure δ. Then

(i) for any vector $\alpha = (\alpha_1, \ldots, \alpha_{k+1})$ of positive components with $\alpha_1 + \ldots + \alpha_{k+1} = 1$, there exists a unique partition R^r similar to ω (i.e., $R^r \in \mathcal{R}_\omega$) with

$$(3) \qquad p_i(r) \equiv p_i(R^r) = \alpha_i, \; i = 1, \ldots, k;$$

(ii) $p_{k+1}(r)$ is a (strictly) increasing and continuous function of r in $(0, \infty)$ and

$$(4) \qquad \lim_{r \to 0} p_{k+1}(r) = \inf_{r > 0} p_{k+1}(r) = \alpha_{k+1}.$$

Proof. Let $p(\delta) = (p_1(\delta), \ldots, p_{k+1}(\delta))$. It may be shown [4] that the set of points $p(\delta)$ for all decision functions δ constitutes a convex and compact subset M of E_{k+1} contained in the unit hypercube K and containing all the corners of K with coordinates adding to 1. The "upper" surface U of M corresponds to the set of admissible procedures, which by Lemma 1(a) are the similar partitions $\mathcal{R}_\omega$ to ω. The line parralel to the k+1-coordinate axis through the point $(\alpha_1, \ldots, \alpha_k, 0)$ intersects U in a single point, namely, the p(r) corresponding to the admissible partition R_r which satisfies (3).

For (ii) note first that $p_{k+1}(r) > \alpha_{k+1}$ for $r > 0$, since otherwise the completely randomized ("guess") procedure δ_o with $p_i(\delta) = \alpha_i$, $i = 1, \ldots, k+1$, would also be admissible.

Now observe that the (unique) admissible partitions $R^r = (R_1^r, \ldots, R_{k+1}^r)$ and $R^{r'} = (R_1^{r'}, \ldots, R_{k+1}^{r'})$ for the σ_r- and $\sigma_{r'}$-classification problems, respectively, which satisfy

(5) $$p_i(r) = p_i(r') = \alpha_i,\ i = 1,\dots,k,$$

have the following relation: the partitioning point $\tau_{r'}$, which defines $R^{r'}$ as the translation $\tau_{r'}$ of ω, lies in R^r_{k+1} whenever $r' < r$. To see this note that if $\tau_{r'}$ were in the complement $\bar{R}^r_{k+1}$ of R^r_{k+1}, then at least one of the $R_i^{r'}$, say $R_{i*}^{r'}$ would be a proper subset of R^r_{i*}. But then, by Lemma 1.7.3. of [2], we would have

$$P[X \in R^r_{i*} \mid \mu = \mu_{i*}(r')] < P[X \in R^r_{i*} \mid \mu = \mu_i(r)] = \alpha_{i*},$$

and hence also

$$P[X \in R^{r'}_{i*} \mid \mu = \mu_{i*}(r')] < \alpha_{i*},\quad \text{for some } i* = 1,\dots,k,$$

which contradicts (5). Therefore $\tau(r')$ lies in the interior of R^r_{k+1}, and , by the same argument of Lemma 1.7.3,

$$p_{k+1}(r') = P[X \in R^{r'}_{k+1} \mid \mu = \mu_{k+1}\ (r')] < P[X \in R^{r'}_{k+1} \mid \mu = \mu_{k+1}(r)]$$

$$< P[X \in R^r_{k+1} \mid \mu = \mu_{k+1}\ (r)] = p_{k+1}(r)$$

Since the continuity of $p_{k+1}(r)$ is an immediate consequence of the continuity of the normal distribution, the monotonicity and continuity of $p_{k+1}(r)$ have been established.

Finally (4) follows from the fact that $p_{k+1}(r)$ is bounded below by α_{k+1}; for when $r = 0$, the vertices $\mu_1(0),\dots,\mu_{k+1}(0)$ coincide with the point μ_0 and clearly the σ_0-classification problem degenerates. However, for the topothetical problem of locating μ into one of the cones $\omega_1,\dots,\omega_{k+1}$ intersecting at μ_0, if $\alpha_i(\delta)$ denotes the minimum probability of correctly taking decision d_i that $\mu \in \omega_i$ when the cecision rule δ is used, then no δ with $\alpha_i(\delta) = \alpha_i$, $i = 1,\dots,k$ can improve upon the completely randomized δ_0 with $\alpha_i(\delta_0) = \delta_i$, $i = 1,\dots,k+1$, i.e., for all these δ, $\alpha_{k+1}\ (\delta) = \alpha_{k+1}$. But $p_i(r) = \alpha_i(R^r)$, $i = 1,\dots,k + 1$

(cf. Lemma 1.7.3), and , therefore, as $r \to 0$

$$p_{k+1}(r) \to \alpha_{k+1} = \inf_{r > 0} p_{k+1}(r),$$

which completes the proof of the lemma.

Lemma 4. For $0 < r \leq 1$, the points τ_r which determine the admissible partitions R^r with $p_i(r) = \alpha_i$; $i = 1,\ldots k$, lie in a compact subset T of E_k.

Proof: Let $0 < \epsilon_i < \alpha_i$ such that $\epsilon_i < \frac{1}{2}$, $i = 1,\ldots,k+1$. Then there are finite negative constants c_{ij} such that for each $i=1,\ldots,k+1$, since X is $N(\mu, I)$,

(6) $$P[\delta_{ij}(X) < c_{ij} \mid \mu = \mu_i] = \epsilon_i, \; j \neq i.$$

Define

$$T_{ij} = \{x \in E_k\colon \delta_{ij}(x) < c_{ij}\}, \; i \neq j.$$

For each r in $(0,1]$, let τ_r denote the point which gives the (optimum) partition R^r of Lemma 3 for which

(7) $$p_i(r) = \alpha_i, \; i = 1,\ldots,k.$$

Note that $\tau(1)$ must be in the complement $\overline{T}_{ij}$ of T_{ij} for each $i \neq j$, since otherwise R_i^1, being a subset of T_{ij} for each $j \neq i$,

$$P[X \in R_i^1 \mid \mu = \mu_i] < P[X \in T_{ij} \mid \mu = \mu_i] = \epsilon_i < \alpha_i.$$

Hence $\tau(1)$ must lie in the intersection

$$T = \bigcap_{i \neq j} \overline{T}_{ij};$$

this is the desired compact set. To see this note that each $\overline{T}_{ij}$ is a half-space determined by the hyperplane H_{ij}; $\delta_{ij}(x) = c_{ij}$ such that by (6) $\mu_0 \in \overline{T}_{ij}$; each H_{ij} intersects every other hyperplane except H_{ji}: $\delta_{ji}(x) = c_{ji}$, and therefore the set T is in general the boundary and the interior of a polyhedron. Hence T is a compact set. If $r < 1$, by Lemma 3, $p_{k+1}(r) > \alpha_{k+1}$ and again τ_r

must lie in T, since, for each i, i=1,...,k+1, if $\tau_r \in T_{ij}$ then by Lemma 1.7.3

$$P[\delta_{ij}(X) < c_{ij} \mid \mu = \mu_i(r)] < P[\delta_{ij}(X) < c_{ij} \mid \mu = \mu_i] = \epsilon_i < \alpha_i.$$

which contradicts (7) (since $P[X \in R_i^r \mid \mu = \mu_i(r)] < P[\delta_{ij}(X) < c_{ij} \mid \mu = \mu_i(r)]$).

Remark. Since the simplex σ is arbitrary, Lemma 4 holds for all bounded positive r.

2. Main results. We are now ready for the proof of the main result of

Theorem 1. Given any vector $\alpha = (\alpha_1, \ldots, \alpha_{k+1})$ of positive components with $\alpha_1 + \ldots + \alpha_{k+1} = 1$, there exists a unique partition $R(\tau(\alpha))$ similar to ω such that

$$P[X \in R_i(\tau(\alpha)) \mid \mu = \mu_0] = \alpha_i,\ i = 1, \ldots, k+1.$$

Proof. The point τ_r of Lemma 4, as a vector function of r, is, like $p_{k+1}(r)$, a continuous function of r, and as r decreases continuously from 1 to 0 the corresponding point τ_r traces a continuous curve (arc) which lies in the compact set T of Lemma 4. Therefore, if $\{r_n\}$ is a decreasing sequence converging to 0 the corresponding sequence of points τ_{r_n} has at least one limit point τ^*, say. Let $R(\tau^*)$ denote the similar partition to ω with vertex τ^*. It follows from Lemma 3 that

$$\begin{aligned} &P[X \in R_{k+1}(\tau^*) \mid \mu = \mu_0] = \lim_{r \to 0} p_{k+1}(r) = \alpha_{k+1} \\ (8)\quad & \\ &P[X \in R_i(\tau^*) \mid \mu = \mu_0] = p_i(r) = \alpha_i,\ i=1,\ldots,k. \end{aligned}$$

Furthermore, there is no other partition similar to ω satisfying (8) since if $\tau' \neq \tau^*$ was another point defining a similar partition $R(\tau') = (R_1(\tau'), \ldots, R_{k+1}(\tau'))$, then at least one of the $R_i(\tau')$, $R_{i*}(\tau')$ say, would be a subset of the $R_{i*}(\tau^*)$ and

$$P[X \in R_{i*}(\tau') \mid \mu = \mu_0] < \alpha_{i*}.$$

Taking $R(\tau(\alpha)) = R(\tau^*)$ completes the proof

From Lemma 1 and Theorem 1, we obtain immediately

Corollary 1. Given any vector $\alpha = (\alpha_1, \ldots, \alpha_{k+1})$ of positive components such that $\sum_{i=1}^{k+1} \alpha_i = 1$, the partition $R(\tau(\alpha))$ which is similar to ω and satisfies

$$P[X \in R_i(\tau(\alpha)) \mid \mu = \mu_0] = \alpha_i, \ i = 1,\ldots,k+1,$$

is an admissible partition for the topothetical problem of locating μ into one of $\omega_1,\ldots,\omega_{k+1}$ on the basis of X from $N(\mu, I)$, and it takes decision d_i correctly with probability at least α_i, $i = 1,\ldots, k+1$. The partition corresponding to

$$\alpha_1 = \alpha_2 = \ldots = \alpha_{k+1} = \frac{1}{k+1}$$

is an admissible minimax for the same problem.

Also, combining the preceding discussion with the observation that there exists one-to-one correspondence between the set of similar partitions $R(\tau)$, and the corresponding probability vectors $\alpha(\tau)$ gives

Corollary 2. Given any vector $\alpha = (\alpha_1,\ldots,\alpha_{k+1})$ of positive components and any partition R of E_k into $k+1$ convex polyhedral cones with the same vertex, there exists a unique point $\mu(\alpha, R)$ such that

$$P[X \in R_i \mid \mu = \mu(\alpha, R)] = \alpha_i, \ i = 1,\ldots,k+1.$$

In addition, the partition R is admissible for the topothetical problem of locating the normal mean μ into one of the $k+1$ cones which constitute the similar partition to R with vertex $\mu(\alpha, R)$; the α_i's determine the minimum probabilities of correct decision.

3. Partitioning E_k into more than $k+1$ regions. So far we have considered the case of $k+1$ convex polyhedral conical regions in E_k. The question now arises as to whether it is possible to partition E_k into more than $k+1$ regions of the same type with preassigned probability contents. We shall indicate by some examples below that, in general, it is no longer possible. This, in a way, reflects the fact that in a classification problem with more than $k+1$ alternatives the class of admissible procedures is no more determined by regions of the type

considered, though still each of the regions is bounded by hyperplanes. (A detailed study of such regions and related topothetical problems will appear in a subsequent paper). Thus, in the case of k + 2 alternatives, the specification of an admissible partition, roughly speaking, requires in general, besides the directions of the bounding hyperplanes, the specification of two points and not one (the translation vector τ) as in the case of k + 1 . Indeed, it should be observed that the limiting argument employed before rests very heavily on the conical shape of the admissible partitions.

Let us consider the case of k + 2 equal components $\alpha_1,\ldots,\alpha_{k+2}$.

Example 1. Let k = 2 and consider 4 angular regions $\omega_1,\ldots,\omega_4$, as in Figure 1(a), specified by two lines intersecting not at right angles. Then there is no partition of the plane similar to $\omega = (\omega_1,\ldots,\omega_4)$ such that the corresponding regions have the same probability content, i.e., one quarter. In fact, any partitioning point (translation) such that the similar regions to ω_1 and ω_3 have the same probability content under the circular normal distribution centered at O, must lie on the bisector of the angle AOC by symmetry of the distribution. Similarly, for ω_2 and ω_4 such a point must lie on the bisector of the angle AOD. Hence the partitioning point would be the vertex O; but then the probability content of ω_i, being proportional to the size of the corresponding angle, is different than a quarter.

A less trivial example is illustrated in Fig. 1(b), where $\widehat{AOB} = \widehat{BOC}$ and the OD is not the bisector of $\widehat{AOC}$ (O the center of the distribution). Clearly any partitioning point that makes the similar regions to ω_1 and ω_2 equiprobable must lie on OB, and then the similar regions to ω_3 and ω_4 have different probability contents.

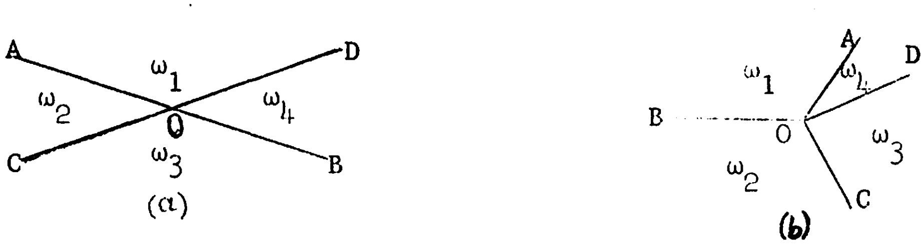

Fig. 1. Nonexistence of four equiprobable angular regions.

Analogous counter-examples can easily be constructed for the case of unequal probabilities α_i, as well as for the case of more than k + 2 cones.

Remark. The nonexistence of similar partitions to a given partition ω with specified probability contents under the k-variate normal distribution when the number m of conical components of ω are more than k + 1 should be expected in view of the fact that a similar partiition to ω is completely specified by its vertex V, and hence there are k unknowns, the coordinates of V, whereas, there are $m-1 > k$ independent equations which, in general, do not have a solution. For $m = k + 1$ however, the k independent equations corresponding to any probability vector $\alpha = (\alpha_1,\ldots,\alpha_{k+1})$, $\alpha_i > 0$, have, by Theorem 1, a unique solution $V = V_\alpha$. Nevertheless, the author has not been able to find a proof of the result by dispensing with the assumption of convexity of the k + 1 cones.

4. Some applications. Several multi-decision problems concerning normal population means may be reduced to the topothetical problem of locating a k-variate normal mean into one of k + 1 convex polyhedral cones. If an indifference region is properly chosen (cf. Lemma 2 and [2]) then a unique admissible minimax partition may be found.

Following are some multiple-decision problems lending themselves to our topothetical approach. The reader may certainly find more such examples.

Example 1. Selecting the largest mean. On the basis of n independent observations $x_\alpha = (x_{1\alpha},\ldots,x_{p\alpha})$, $\alpha = 1,\ldots, n$ on X distributed according to $N(\mu, \Sigma)$, choose the largest component of $\mu = (\mu_1,\ldots,\mu_p)$ assuming that the covariance matrix Σ is known. By invariance considerations the selection procedures may be based on a maximal invariant $y = (\bar{x}_2 - \bar{x}_1,\ldots,\bar{x}_p - \bar{x}_1)$, say, where

$\bar{x}_i = \frac{1}{n} \sum_{i=1}^{n} x_{i\alpha}$, $i = 1,\ldots,p$. Clearly y is $N(\delta, \Sigma^*)$, where $\delta = (\mu_2-\mu_1,\ldots,\mu_p-\mu_1)$ and Σ^* is a known $(p-1) \times (p-1)$ positive definite matrix. It is seen that the decision d_i, that μ_i is the largest, is appropriate when δ lies in a convex (p-1) -dimentional polyhedral cone in the (p-1)-space of δ. The same reduction holds when we have n_i observations on the i-th component of X. In the statistical literature the case of $\Sigma = \sigma^2 I$ has been studied quite extensively both when σ^2 is known or unknown.

Example 2. A slippage problem. This may be obtained as a special case of Example 1 if we assume that all the components of μ are equal except one which is larger (slips to the right by some positive amount $\Delta > 0$). Here we do not allow the possibility of all the μ_i being equal. The decision d_i that μ_i slipped corresponds to points δ on a ray through the origin. It follows that if Δ is bounded away from zero, i.e., $\Delta \geqq \epsilon > 0$, then by Lemma 2 (b) there exists a unique admissible minimax procedure.

REFERENCES

[1] Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis. Wiley, New York.

[2] Cacoullos, T. (1962). Comparing Mahalanobis Distances. Columbia Ph.D. Dissertation, to appear in Sanklyā

[3] Dubins, L. E. and Spanier,E. H. (1961). How to cut a cake fairly. Amer. Math. Monthly 68 1-17.

[4] Dvoretzky, A.,Wald, A., and Wolfowitz, J. (1951). Elimination of randomization in certain statistical decision procedures and zero-sum two-person games. Ann. Math. Statist. 22 1-21.

[5] Wald, A. (1950). Statistical Decision Functions. Wiley, New York.